THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 7, September, 1973
Printed in U.S.A.

A Compromise Equalizer Design Incorporating Performance Invariance

By F. J. BROPHY, G. J. FOSCHINI, and R. D. GITLIN

(Manuscript received February 20, 1973)

We give a solution to the problem of designing a fixed compromise equalizer for use in transmission systems involving an ensemble of random channels. The signal and noise spectra, along with the second-order statistics of the channel ensemble, are used to find the equalizer characteristic that minimizes the mean-square distortion between the equalizer output and a scaled version of the transmitter output. The key departure from previous work is that the criterion better captures practical performance invariance; specifically, the cost function incorporates the insensitivity of a well-designed demodulator to any amplitude scaling or time delay introduced by a particular channel. After demonstrating that the optimum equalizer shape is related to the principal eigenfunction of a normalized channel correlation function, we consider several special cases that give further insight into the properties of the solution. We find that the equalizer amplitude is attenuated over those frequencies where the signal-to-noise or signal-to-channel-variance ratios are small. The analysis confirms the standard engineering practice of inverting the average channel in the absence of noise and when the variance of the channel characteristics is small.

I. INTRODUCTION

A fixed compromise equalizer is frequently employed in data transmission systems to compensate for linear distortion introduced by a channel drawn from a random ensemble.[1] Typically, compromise equalizers find application in systems which, because of economic or other considerations, do not use an adaptive equalizer. It is possible that, even when adaptive equalization is used to compensate for a particular channel characteristic, one might use a compromise equalizer to provide a good initial channel and thereby reduce the receiver

adaptation time. As its name suggests, the equalizer is a *fixed* linear filter that effects a compromise by compensating for an "average" channel. We propose a procedure that uses the statistics of the channel ensemble, the modulated signal, and the additive noise at the demodulator to design a filter that minimizes a performance measure appropriate for most transmission systems. The performance measure is an adaptively scaled mean-square error. For example, it is particularly well suited for use in a data transmission system and results in a filter that is significantly different from that obtained by directly minimizing the mean-square error.[2]

In order to avoid being restricted by nonlinear demodulation techniques (e.g., those used in FSK or DPSK data systems) and to be able to accommodate asynchronous signalling, the equalizer operates directly on the received passband signal and is thus a channel, rather than a synchronous, equalizer. Our performance criterion is the continuous time mean-square error between the equalizer output and an adaptively scaled version of the transmitter output. This scaling needs to be done only once in the design of the filter and not each time a new channel is dialed up. The scaling is such that the filter is invariant to any amplitude scaling, sign inversion, or time delay encountered in transmission over a particular channel; this type of invariance is appropriate for a compromise equalizer used in most transmission systems, since amplitude scaling and a fixed time delay will not be "seen" by a well-designed demodulator.

Since the criterion is quadratic in nature, the optimum filter response is determined by the second-order statistics of the signal, noise, and channel ensemble. The best filter shape is shown to be the principal eigenfunction of an integral operator whose kernel is a weighted channel correlation function. An explicit design procedure is described in the text, and several interesting questions are discussed as well. Some of these questions are:

(*i*) How different is the compromise equalizer from the inverse of the average channel characteristic?

(*ii*) Suppose there is no amplitude distortion but only delay distortion. What is the nature of the compensation? Is there amplitude as well as delay compensation?

(*iii*) How sensitive is the filter design to different signal spectra?

In Section II the compromise equalization problem is formulated, and the distortion measure is discussed in considerable detail. The determination of the optimum filter is described in Section III, and an example is provided in Section IV illustrating the design technique.

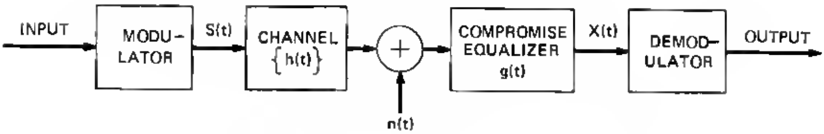

Fig. 1—Transmission system with a compromise equalizer.

II. PROBLEM FORMULATION

2.1 *System Configuration*

We begin by specifying the framework of our discussion. To accommodate a system employing nonlinear demodulation, tbe compromise equalizer is assumed to be placed in the passband. Since we also want to provide a setting general enough to include asynchronous transmission, no restriction is placed on the form of the filter other tban requirement of a finite energy impulse response. As shown in Fig. 1, $s(t)$ denotes the modulated signal, $n(t)$ the received additive noise, $x(t)$ the compromise equalizer output, and $g(t)$ the equalizer impulse response (wbich we ultimately seek to specify). The ensemble of channel impulse responses, $\{h(t)\}$, as well as $n(t)$ and $s(t)$, are all assumed to be continuous in quadratic mean and statistically independent of eacb otber. Botb $s(t)$ and $n(t)$ will be taken to be zero mean, to be wide-sense stationary with finite power and to possess power spectral densities $S(\omega)$ and $N(\omega)$, respectively. To make the subsequent analysis precise, we assume $h(t)$ has bounded absolutely integrable sample paths (with probability one); thus, the ensemble of channel frequency characteristics $\{H(\omega)\}$ is well defined.†

Our problem is first to select an appropriate optimization criterion and then to choose the compromise equalizer that minimizes this measure. To avoid the specifics of particular demodulators, we will be concerned witb preserving the fidelity between $s(t)$ and $x(t)$. A natural first choice for a distortion measure is the continuous-time mean-square error

$$\mathcal{E}_0 = \langle [x(t) - s(t-\tau)]^2 \rangle, \tag{1}$$

where $\langle\ \rangle$ denotes the average with respect to the signal, noise, and cbannel ensemble, and τ is an arbitrary delay. Using this measure, Maurer and Franks[2] found that tbe optimum filter is given by

$$G_o(\omega) = \frac{S(\omega)\langle H(\omega)\rangle^* e^{-j\omega\tau}}{S(\omega)\langle |H(\omega)|^2\rangle + N(\omega)}, \tag{2}$$

† The process $\{H(\omega)\}$ is assumed to have a square integrable covariance, and the variance is assumed to be nonzero for the frequency range of interest.

where the asterisk denotes the complex conjugate and $H(\omega)$ is a member of the channel ensemble. It is easy to see that $G_o(\omega)$, given by (2), is not just the Wiener filter for the average channel, since the mean-square channel dispersion $\langle |H(\omega)|^2\rangle$, rather than $|\langle H(\omega)\rangle|^2$, modifies the filter characteristic. We note that, even as $N(\omega) \to 0$, the filter does not generally invert the average channel. This will be the case when the channel variance is extremely small so that $\langle |H(\omega)|^2\rangle \approx |\langle H(\omega)\rangle|^2$; this would occur, for instance, when the ensemble consists of only one characteristic.

Care must be exercised in determining how to use the design technique that results in the filter given by (2). For example, suppose the individual transfer characteristics are greatly varied in the degree of attenuation they impart across the band of interest. It can follow (depending on how probabilities are attached) that the low-loss channels determine the character of the averages in (2)—the "whomper" effect. Hence, linear distortion in the lossy channels may not be suitably equalized. However, if linear distortion is uniformly the dominant impairment, the noise level can be set to zero and the "whomper" effect easily eliminated by preparing the data by normalizing the channels prior to averaging. For example, the normalization can be accomplished by individually scaling the characteristics so they have the same energy in response to some pulse or so they have the same gain at some central frequency. Even if linear distortion is not the uniformly dominant impairment, there may be applications where the noise is set to zero and a normalization of the channels made prior to computing the optimum equalizer. In such a procedure, one is trading noise immunity for immunity to linear distortion. Another case of importance in applications is when the channels have the same amplitude characteristic but different phase characteristics; here, of course, the "whomper" effect is nonexistent. The considerations of this paragraph will also apply to the design technique we shall develop in the following sections.

While the $\mathcal{E}_0$ criterion provides an interesting and tractable formulation, it has the shortcoming that it understates the capability of most demodulators. A striking example of this occurs if the ensemble $\{h(t)\}$ has zero mean. The zero mean is reasonable for systems subject to occasional "phase hits," the mathematical implication being that, if $h(t)$ is a possible channel, then $-h(t)$ is just as probable. Then $G_o(\omega) \equiv 0$. Yet, in practice, one would expect to do better than "pull out the plug and go home."

At the other extreme we could use an information theoretic type

criterion. For example, we could choose $G(\omega)$ to maximize the average mutual information between $x(t)$ and $s(t)$. This criterion is mathematically tractable; in fact, the solution is trivial—$G(\omega)$ can be any characteristic which is nonzero over the same frequency range as $S(\omega)$.† The shortcoming of this criterion is (as the solution suggests) that it overestimates the demodulation capability.

As the state of the art in demodulation advances, we would anticipate an evolution of criteria away from the mean-square error toward the information theoretic. Clearly, a good criterion should give the demodulator credit for what it can realistically accomplish and at the same time pose a tractable optimization problem for determining $G(\omega)$.

With this motivation, we propose a mean-square-error criterion which reflects the fact that the performance of a well-designed demodulator is insensitive to scaling of the input (the automatic gain control feature), an input sign change (the information is generally differentially encoded), and, of course, a time delay. Under such a criterion, the compromise equalizer will no longer be implicitly constrained by attempting to faithfully reproduce the modulated signal.

2.2 *An Adaptively Scaled Mean-Square Error*

Based upon the above discussion, we consider the following adaptively scaled mean-square error

$$\mathcal{E} = \min_{A,B} \langle [x(t) - As(t - B)]^2 \rangle_{s,n}, \tag{3}$$

where A and B are real numbers and the averaging is over the signal and noise statistics with both the channel and equalizer held fixed. The criterion is meaningful under the assumption that the signal-to-noise ratio at the receiver does not change appreciably from channel to channel. We stress that the considerations mentioned in the third paragraph of Section 2.1 apply here, as well. The optimum equalizer is obtained by averaging $\mathcal{E}$ over the channel ensemble and then minimizing the result with respect to the equalizer transfer function $G(\omega)$, subject to a power constraint on the demodulator input. The quantities A and B, which are channel-dependent, provide an adaptively scaled reference signal $As(t - B)$. The reference is adaptive in that, for each realization of the channel, A and B are chosen to minimize $\mathcal{E}$. Notice, for example, that, if a particular channel introduces a sign inversion, then $A = -1$ will remove this effect. When a channel

† This comes about because no information is lost when the signal is subject to a reversible operation (such as a channel).

has a more complicated phase and/or gain characteristic, it is no longer apparent what the optimal value of A should be; however, we shall shortly see that the minimizing value of A can be determined analytically. We shall also consider the determination of B. Thus, the filter will not expend any of its degrees of freedom by attempting to compensate for a sign inversion, amplitude scaling, or time delay introduced during transmission. It should be clear that, since A and B depend on the channel characteristics, they are random variables. Simply put, the criterion given by (3) forces the equalizer to minimize only the portion of the output signal that does not look like a scaled or delayed version of the transmitted signal.

We now consider the properties of the adaptively scaled mean-square error. We begin by letting

$$I = \langle [x(t) - As(t-B)]^2 \rangle_{s,n}, \tag{4}$$

i.e.,

$$\mathcal{E} = \min_{A,B} I.$$

Carrying out the indicated average gives

$$I = \int_{-\infty}^{\infty} \{S(\omega)[|F(\omega)|^2 - 2Ae^{j\omega B}F(\omega) + A^2] + |G(\omega)|^2 N(\omega)\} \frac{d\omega}{2\pi}, \tag{5}$$

where $S(\omega)$ is the power spectral density of $s(t)$ and $F(\omega)$ is the product of $G(\omega)$ and $H(\omega)$. (Notice I does not change when $F(\omega)$ is replaced by $Re\{F(\omega)\}$.) To find the minimum of I with respect to A and B, we set to zero the partial derivatives of I with respect to these variables and find that

$$\frac{\partial I}{\partial A} = -2\int_{-\infty}^{\infty} S(\omega)e^{j\omega B}F(\omega)\frac{d\omega}{2\pi} + 2A\int_{\infty}^{\infty} S(\omega)\frac{d\omega}{2\pi} = 0 \tag{6a}$$

$$\frac{\partial I}{\partial B} = -2\int_{-\infty}^{\infty} j\omega S(\omega)e^{j\omega B}F(\omega)\frac{d\omega}{2\pi} = 0. \tag{6b}$$

From (6a) we have A_{opt} given by the correlation ratio

$$A_{\text{opt}} = \frac{\displaystyle\int_{-\infty}^{\infty} S(\omega)e^{j\omega B}F(\omega)\frac{d\omega}{2\pi}}{\displaystyle\int_{-\infty}^{\infty} S(\omega)\frac{d\omega}{2\pi}} = \frac{\langle s(t-B)x(t)\rangle}{\langle s^2(t-B)\rangle}. \tag{7}$$

The interpretation of B_{opt} is facilitated by letting

$$y(t) \equiv \int_{-\infty}^{\infty} j\omega F(\omega)S(\omega)e^{j\omega t}\frac{d\omega}{2\pi}, \tag{8}$$

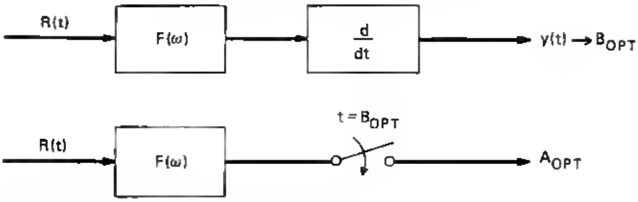

Fig. 2—Interpretation of optimum delay (B_{opt}) and amplitude (A_{opt}) scaling. B_{opt} is one of the instants when $y(t)$ is zero.

which is recognized as the response of a linear filter with transfer function $j\omega F(\omega)$ [i.e., $F(\omega)$ followed hy a differentiator] to the input $R(t)$, where $R(t)$ is the (inverse) Fourier transform of $S(\omega)$ and is the signal correlation function. Comparing (6h) and (8), we see that B_{opt} is one of the instants when $y(t)$ is zero. We illustrate this interpretation of B_{opt} in Fig. 2, where we also indicate how A_{opt} may he ohtained in a similar manner. Determining B_{opt} is a very difficult problem, since it is tantamount to asking for the zero crossing of a signal from knowledge of its Fourier transform. In order to proceed further, we will approximate B_{opt} hy the delay at midhand, which we conveniently take to he zero for each channel.

Using the value of A_{opt} given hy (7) and setting $B = 0$, we have

$$\mathcal{E} = \int_{-\infty}^{\infty} |G(\omega)|^2[S(\omega)|H(\omega)|^2 + N(\omega)]\frac{d\omega}{2\pi} - \frac{1}{\alpha}\left|\int_{-\infty}^{\infty} S(\omega)G(\omega)H(\omega)\frac{d\omega}{2\pi}\right|^2, \quad (9)$$

where

$$\alpha = \int_{-\infty}^{\infty} S(\omega)\frac{d\omega}{2\pi}.$$

The optimum filter is ohtained hy first averaging (9) with respect to the ensemble of channels and then minimizing this average with respect to $G(\omega)$.† Before doing this, we give geometric interpretations to hoth the criterion and the optimum filter.

2.3 *A Geometric Interpretation of the Problem*

Finding a geometric framework in which to view the optimization prohlem will he quite helpful in understanding the nature of the solution. To put the prohlem at hand in such a setting, we introduce

† The minimization is done subject to a constraint on the average output power.

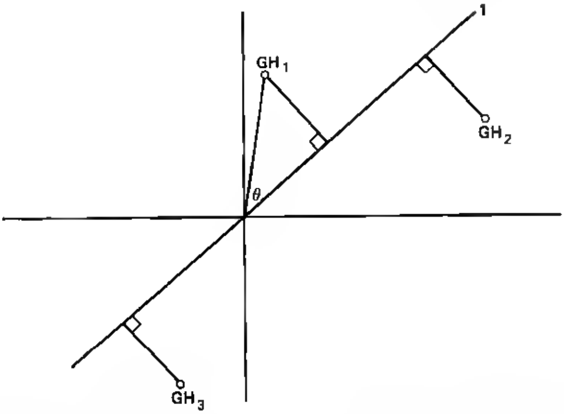

Fig. 3—Geometric interpretation of criterion. The adaptively scaled mean-square error may be interpreted as the average squared distance from the point **GH** (which represents an equalized channel) to the ray **1** (which represents a distortionless channel).

the signal-weighted inner product

$$(\mathbf{U}, \mathbf{V}) \equiv \int_{-\infty}^{\infty} U(\omega) V^*(\omega) S(\omega) \frac{d\omega}{2\pi}, \tag{10}$$

where the vectors **U** and **V** represent the functions $U(\omega)$ and $V(\omega)$ respectively, and the vector **1** will correspond to the real function of unit amplitude. Suppose that the noise is set to zero; then, in terms of the above notation, we can write†

$$\mathcal{E} = \|\mathbf{GH}\|^2 - \frac{1}{\alpha}(\mathbf{GH}, \mathbf{1})^2, \tag{11}$$

where the vector **GH** corresponds to the function $G(\omega)H(\omega)$. For convenience we set the signal power equal to unity, i.e., $\alpha = 1$, and apply the Schwarz inequality, which gives

$$\mathcal{E} = \|\mathbf{GH}\|^2 - (\mathbf{GH}, 1)^2 \geqq \|\mathbf{GH}\|^2 - \|\mathbf{GH}\|^2 \cdot \|\mathbf{1}\|^2 = 0, \tag{12}$$

where the lower hound is achieved when **GH** is proportional to **1**.‡ From (11) we have

$$\begin{aligned} \mathcal{E} &= (\mathbf{GH}, \mathbf{GH}) - (\mathbf{GH}, \mathbf{1})^2 = (\mathbf{GH}, \mathbf{GH})\left[1 - \frac{(\mathbf{GH}, \mathbf{1})^2}{(\mathbf{GH}, \mathbf{GH})}\right] \\ &= \|\mathbf{GH}\|^2[1 - \cos^2\theta] = [\|\mathbf{GH}\| \sin\theta]^2, \end{aligned} \tag{13}$$

† The norm of the vector **U**, denoted by $\|\mathbf{U}\|$, is given by $(\mathbf{U}, \mathbf{U})^{\frac{1}{2}}$.

‡ Thus, if we have only one channel characteristic and no noise, the equalizer will invert the channel.

where θ is the angle between $\mathbf{GH}$ and $\mathbf{1}$. Equation (13) provides a very useful interpretation of the error $\mathcal{E}$. As shown in Fig. 3, $\|\mathbf{GH}\| \sin \theta$ is the distance from the vector $\mathbf{GH}$ to the ray colinear with the vector $\mathbf{1}$. Hence, the average of $\mathcal{E}$ with respect to the channel ensemble, which we denote by $\langle \mathcal{E} \rangle_{\mathbf{H}}$, is the average squared distance from $\mathbf{GH}$ to the ray $\mathbf{1}$. Thus, for a given channel-equalizer constellation, $\{\mathbf{GH}_i\}$, the equalizer is chosen to minimize the dispersion about the ray $\mathbf{1}$.

In order to get a feeling for the capability of the equalizer to modify the channel constellation $\{\mathbf{H}_i\}$, we associate with $G(\omega)$ a linear operator $\mathcal{G}$ that maps a particular channel $\mathbf{H}_i$ into $\mathbf{GH}_i$; we write this operation symbolically as

$$\mathcal{G}\colon\ \mathbf{H}_i \rightarrow \mathbf{GH}_i. \tag{14}$$

The equalizer operator, $\mathcal{G}$, is called diagonal since it modifies $H(\omega)$ in a pointwise fashion to produce $G(\omega)H(\omega)$. Suppose, for the purpose of illustration, we relax our hypothesis and assume that the channel ensemble has energy only at two values of ω, ω_1 and ω_2. In this case, the operator $\mathcal{G}$ is particularly simple since the point $\mathbf{H} = (h_1, h_2)$ is mapped into the point $\mathbf{GH} = (g_1h_1, g_2h_2)$, where $h_i \equiv H(\omega_i)$ and $g_i \equiv G(\omega_i)$. The locus of points (g_1h_1, g_2h_2), subject to the average power constraint on the demodulator input $x(t)$

$$g_1^2k_1 + g_2^2k_2 = 1 \qquad (k_i \triangleq E\{|H(\omega_i)|^2S(\omega_i)\}), \tag{15}$$

describes the manner in which the equalizer redistributes the channel ensemble so as to minimize the average squared distance to the ray $\mathbf{1}$.

By noting that the point (g_1h_1, g_2h_2) satisfies the relation

$$\frac{(g_1h_1)^2k_1}{(h_1)^2} + \frac{(g_2h_2)^2k_2}{(h_2)^2} = 1, \tag{16}$$

we see that the locus of the points (g_1h_1, g_2h_2) subject to (15) is an ellipse centered around the origin whose major and minor axes are parallel to the x–y axes and are of length $|h_1/k_1^{\frac{1}{2}}|$ and $|h_2/k_2^{\frac{1}{2}}|$. Thus, the effect of the equalizer on the channel array, as shown in Fig. 4, is to allow each equalized channel, $\mathbf{GH}_i = \{gh_1^{(i)}, gh_2^{(i)}\}$, to move on an elliptical surface. The nature of the compromise by which the optimum filter shape is chosen should now be clear. Each channel wants to move as close to the unit ray as possible—this will usually lead to conflicting requirements, i.e., as one channel is brought closer to the unit ray, other channels will move further away from this ray. For a multiplicity of channels there will be an "elliptical flow" along the respective ellipses which terminates when the best filter shape has been found.

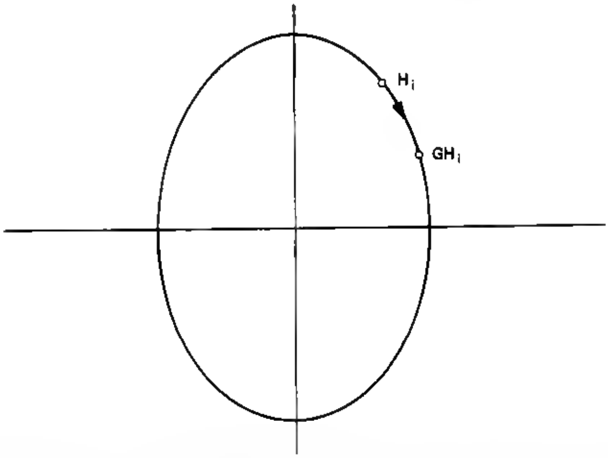

Fig. 4—Geometric interpretation of tbe effect of the compromise equalizer on the channel ensemble. A channel, represented by the point $\mathbf{H}_j$, is modified by the equalizer to give tbe equalized cbannel $\mathbf{GH}_j$.

The ahove interpretation would of course be valid in a Hilhert space of dimension large enough to accurately represent the function $H(\omega)$.

III. DETERMINING THE OPTIMUM EQUALIZER CHARACTERISTIC

3.1 *Analytic Solution*

Having developed a geometric representation as an aid to understanding the equalizer design prohlem, we are now in a position to explicitly determine the best filter shape. Returning to (9) and averaging with respect to the channel ensemhle, we have

$$\langle \mathcal{E} \rangle_H = \int_{-\infty}^{\infty} |G(\omega)|^2 [S(\omega) H_{\text{rms}}^2(\omega) + N(\omega)] \frac{d\omega}{2\pi} - \frac{1}{\alpha} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\omega) S(\nu) G^*(\omega) G(\nu) \langle H^*(\omega) H(\nu) \rangle \frac{d\omega}{2\pi} \frac{d\nu}{2\pi}, \quad (17)$$

where

$$H_{\text{rms}}(\omega) \equiv \sqrt{\langle |H(\omega)|^2 \rangle}. \quad (18)$$

If we let

$$Q(\omega) = \sqrt{S(\omega) H_{\text{rms}}^2(\omega) + N(\omega)}\, G(\omega) \quad (19)^\dagger$$

and introduce the weighted channel covariance kernel

$$K(\omega, \nu) = \frac{1}{\alpha} \frac{S(\omega) S(\nu) \langle H^*(\omega) H(\nu) \rangle}{\sqrt{[S(\omega) H_{\text{rms}}^2(\omega) + N(\omega)][S(\nu) H_{\text{rms}}^2(\nu) + N(\nu)]}}, \quad (20)$$

† Note that $|Q(\omega)|^2$ is the power spectral density of the equalizer output for the "rms cbannel."

then our cost function can be rewritten in the convenient form

$$\langle \mathcal{E} \rangle_H = \int_{-\infty}^{\infty} Q(\omega) Q^*(\omega) \frac{d\omega}{2\pi} - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(\omega, \nu) Q^*(\omega) Q(\nu) \frac{d\omega}{2\pi} \frac{d\nu}{2\pi}. \quad (21)$$

Introducing the Hermitian integral operator $\mathcal{K}$, whose kernel is $K(\omega, \nu)$, permits us to write the criterion as the quadratic form

$$\langle \mathcal{E} \rangle_H = (\mathbf{Q}, \mathbf{Q}) - (\mathcal{K}\mathbf{Q}, \mathbf{Q}), \quad (22)$$

where the inner product does not include the signal-spectrum weighting introduced in (10).† We now wish to minimize (22) with respect to Q, subject to an appropriate constraint. Referring to Fig. 1, we see that the average power present at the demodulator input is

$$P = \int_{-\infty}^{\infty} [S(\omega) H_{\mathrm{rms}}^2(\omega) + N(\omega)] |G(\omega)|^2 \frac{d\omega}{2\pi}.$$

Thus, a natural constraint is to require that the power be constant. In terms of $Q(\omega)$, this constraint takes the form

$$(\mathbf{Q}, \mathbf{Q}) = P, \quad (23)$$

and the optimization problem consists of minimizing the positive definite form (22) subject to (23). The solution, which we denote by Q_{opt}, is easily obtained[3] by using a Lagrange multiplier and is recognized as the *principal eigenfunction*‡ of the operator $\mathcal{K}$. The best filter shape, $G_{\mathrm{opt}}(\omega)$, is obtained from $Q_{\mathrm{opt}}(\omega)$ by using (19), and can he regarded as the first term in a Karhunen-Loève representation[4] of the random process

$$\frac{S(\omega)}{\sqrt{S(\omega) H_{\mathrm{rms}}^2(\omega) + N(\omega)}} H(\omega). \quad (24)$$

The residual value of the criterion, evaluated when $\mathbf{Q} = \mathbf{Q}_{\mathrm{opt}}$, is given by

$$\langle \mathcal{E} \rangle_{\mathrm{opt}} = (\mathbf{Q}_{\mathrm{opt}}, \mathbf{Q}_{\mathrm{opt}})[1 - \lambda] = (1 - \lambda) P, \quad (25)$$

where λ is the maximum eigenvalue and is a measure of how effectively the compromise equalizer performs for various channel ensemhles and various signal and noise spectra. The degree to which the random process given by (24) is described by the first term in a Karhunen-Loève expansion will, of course, determine the equalizer performance.

† In the sequel, the inner product between the vector **U** and **V** will be taken to be $(\mathbf{U}, \mathbf{V}) = \int_{-\infty}^{\infty} U(\omega) V^*(\omega) (d\omega/2\pi)$.

‡ That is, the eigenfunction corresponding to the maximum eigenvalue (it is well known that the eigenvalues of a Hermitian operator are real).

In order to get more insight into the nature of $G_{\text{opt}}(\omega)$ we will, in the next paragraph, consider the form of the optimum filter under some special circumstances. This discussion, along with the examples treated in the next section, will reveal some properties of the optimum filter shape.

3.2 *Special Cases*

3.2.1 *One Channel*

Suppose the channel ensemble consists of only one member,† $H(\omega)$. The principal eigenfunction of $\mathcal{K}$ is easily determined to be

$$Q_{\text{opt}}(\omega) = \frac{S(\omega)}{\sqrt{S(\omega)\,|H(\omega)|^2 + N(\omega)}}\,H^*(\omega). \tag{26}$$

Thus, the best filter shape is

$$G_{\text{opt}}(\omega) = \frac{S(\omega)H^*(\omega)}{S(\omega)\,|H(\omega)|^2 + N(\omega)}. \tag{27}$$

The filter given by (27) is just the well-known Wiener filter for estimating a random signal that has been passed through a nonrandom channel and then corrupted by additive noise. The amplitude characteristic of the filter, which is given by

$$\frac{S(\omega)\,|H(\omega)|}{S(\omega)\,|H(\omega)|^2 + N(\omega)}, \tag{28}$$

provides *noise rejection* at those frequencies where $S(\omega)$ is small relative to $N(\omega)$. Thus, as the spectrum of the modulating signal is changed, the noise-rejecting regions of the filter will be altered. It is worth noting that, if the channel has only phase distortion, amplitude as well as phase compensation is required.‡ It should also be noted that the amplitude response of the filter will be greatly attenuated at those frequencies where $|H(\omega)|^2$ dominates $N(\omega)/S(\omega)$; this phenomenon, which we call *channel-variance rejection*, is observed in practice for an arbitrary channel ensemble at the frequencies where the variance of $H(\omega)$, defined by

$$\text{Var}[H(\omega)] \equiv \langle |H(\omega)|^2 \rangle - |\langle H(\omega)\rangle|^2, \tag{29}$$

becomes large. Since the phase of $G_{\text{opt}}(\omega)$, for the simple case of one channel, is just the negative of the channel phase, we have, under the further specialization of vanishingly small noise power, the not-

† This assumption gives good insight when there is small dispersion about the average channel, i.e., the channels all look pretty much alike.

‡ This, along with the preceding sentence, provides a partial answer to the second and third questions posed in the introduction.

surprising result that the equalizer should invert the channel. On the other hand, as the noise becomes dominant, the solution is observed to approach a filter matched to the corresponding average channel characteristic and signal spectrum.

3.2.2 *Deterministic Amplitude Distortion, Random Delay Distortion, and No Noise*

Suppose the members of the channel ensemble can be written as

$$H(\omega) = a(\omega)e^{j\theta(\omega)}, \tag{30}$$

where the amplitude response $a(\omega)$ does not vary from channel to channel, and the phase response $\theta(\omega)$ is randomly selected.

We consider first the kernel $K(\, , \,)$ when there is only delay distortion (i.e., $a(\omega) = 1$) and no noise. Since $H_{\text{rms}}(\omega)$ is unity, we have

$$K(\omega, \nu) = \frac{\sqrt{S(\omega)S(\nu)}}{\alpha} \langle e^{j\theta(\omega) - j\theta(\nu)} \rangle, \tag{31}$$

and the filter is given by

$$G_{\text{DELAY}}(\omega) = \frac{Q_{\text{opt}}(\omega)}{\sqrt{S(\omega)}}. \tag{32}$$

If we no longer restrict $a(\omega)$ to be unity, the kernel is still given by (31) while the filter is seen to be

$$G(\omega) = \frac{1}{a(\omega)} G_{\text{DELAY}}(\omega). \tag{33}$$

The above indicates that, in the presence of deterministic amplitude distortion and random phase distortion, the compensation is decoupled in the sense that the filter shape is a cascade of the compensation for the component distortions. While this sort of decoupling is generally not the case, we have found in the example described in the next section that the *phase* characteristic of the equalizer is rather insensitive to changes in amplitude distortion and noise level.

3.2.3 *Small Variation in the Amplitude and Phase Characteristics*

We now wish to give a suggestive description of the optimum filter when the channel characteristics have small variation about the average characteristic. To this end, let us denote a typical characteristic by

$$\begin{aligned} H(\omega) &= a(\omega)e^{j\theta(\omega)} \\ &= \bar{a}(\omega)b(\omega)e^{j[\bar{\theta}(\omega)+\phi(\omega)]}, \end{aligned} \tag{34}$$

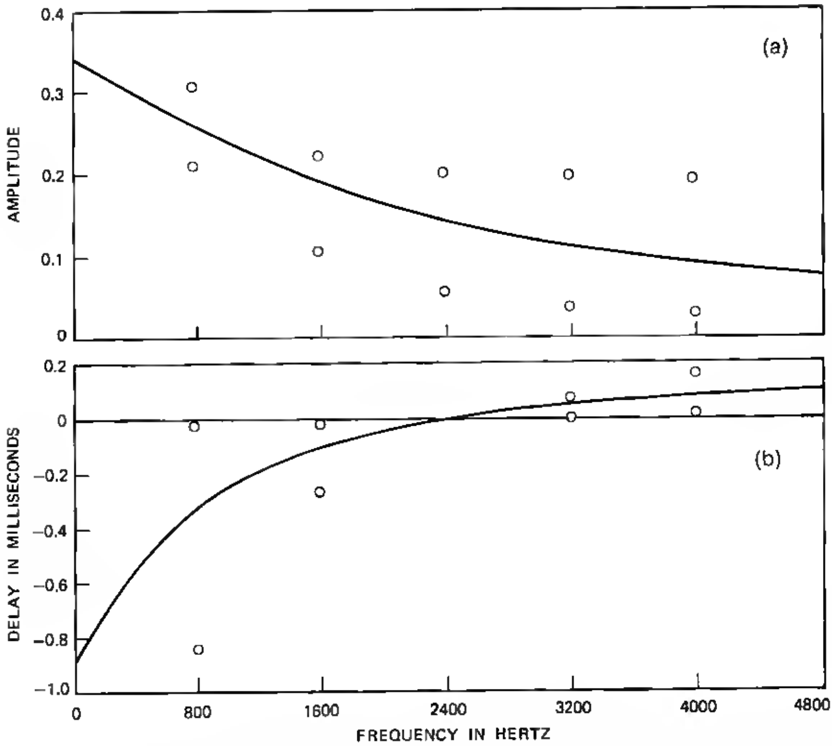

Fig. 5—(a) amplitude and (b) delay characteristics for average channel. The worst case amplitude and phase variation are indicated by the isolated circles.

where $\bar{a}(\omega)$ and $\bar{\theta}(\omega)$ are the average channel amplitude and phase, respectively, and $b(\omega)$ and $\phi(\omega)$ represent small (random) perturbations about these quantities. We first write $H(\omega)$ as

$$H(\omega) = \bar{a}(\omega)e^{j\bar{\theta}(\omega)}e^{\ell n b(\omega)+j\phi(\omega)} \tag{34a}$$

$$= \bar{H}(\omega)e^{q(\omega)}, \tag{34b}$$

where we have let

$$\bar{H}(\omega) = \bar{a}(\omega)e^{j\bar{\theta}(\omega)} \tag{35}$$

$$q(\omega) = \ell n b(\omega) + j\phi(\omega), \tag{36}$$

and we note that $\bar{H}(\omega)$ is composed of the average amplitude and phase of the ensemble. Since $|b(\omega)| \approx 1$ and $|\phi(\omega)| \approx 0$, we see that $|q(\omega)| \approx 0$; thus, retaining only the leading term in the series expansion of $e^{q(\omega)}$ gives

$$e^{q(\omega)} \approx 1 + q(\omega). \tag{37}$$

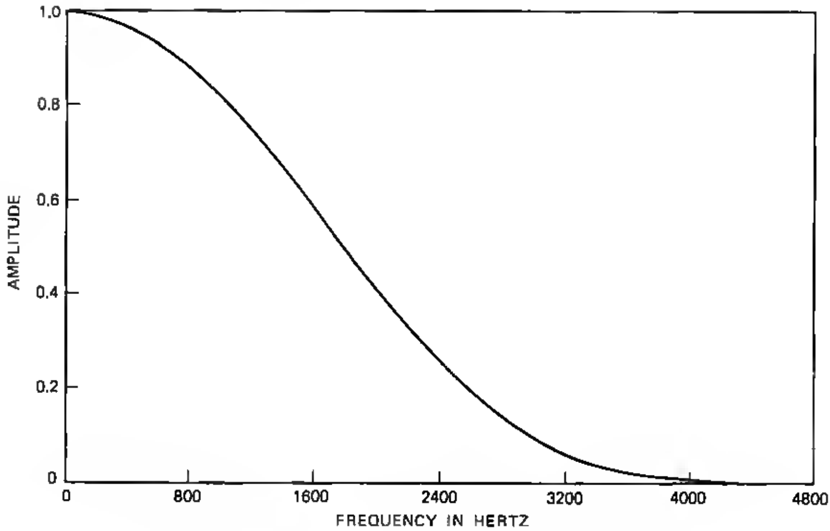

Fig. 6—Raised cosine power spectral density (rolloff = 1).

We are now in a position to evaluate the kernel $K(\omega, \nu)$ under the above assumptions. Using (37) we have

$$\begin{aligned} \langle H^*(\omega)H(\nu)\rangle &= \bar{H}^*(\omega)\bar{H}(\nu)\langle e^{q^*(\omega)}e^{q(\nu)}\rangle \\ &\approx \bar{H}^*(\omega)\bar{H}(\nu)\langle 1 + q^*(\omega) + q(\nu)\rangle, \end{aligned} \tag{38}$$

where we have kept only first-order terms in tbe expansion of $e^{q^*(\omega)}e^{q(\nu)}$. If we assume that the perturbations $\ell nb(\omega)$ and $\phi(\omega)$ have zero average value, then (38) separates, reducing to

$$\langle H^*(\omega)H(\nu)\rangle = \bar{H}^*(\omega)\bar{H}(\nu). \tag{39}$$

By a similar argument, the denominator of the kernel is found to be

$$\sqrt{S(\omega)\,|\bar{H}(\omega)|^2 + N(\omega)} \times \sqrt{S(\nu)\,|\bar{H}(\nu)|^2 + N(\nu)}, \tag{40}$$

and combining (39) and (40) gives the separable kernel

$$K(\omega, \nu) = \frac{S(\omega)S(\nu)\bar{H}^*(\omega)\bar{H}(\nu)}{\sqrt{S(\omega)\,|\bar{H}(\omega)|^2 + N(\omega)}\sqrt{S(\nu)\,|\bar{H}(\nu)|^2 + N(\nu)}}. \tag{41}$$

The equalizer sbape is then given by

$$G_{\text{opt}}(\omega) = \frac{S(\omega)\bar{H}^*(\omega)}{S(\omega)\,|\bar{H}(\omega)|^2 + N(\omega)}, \tag{42}$$

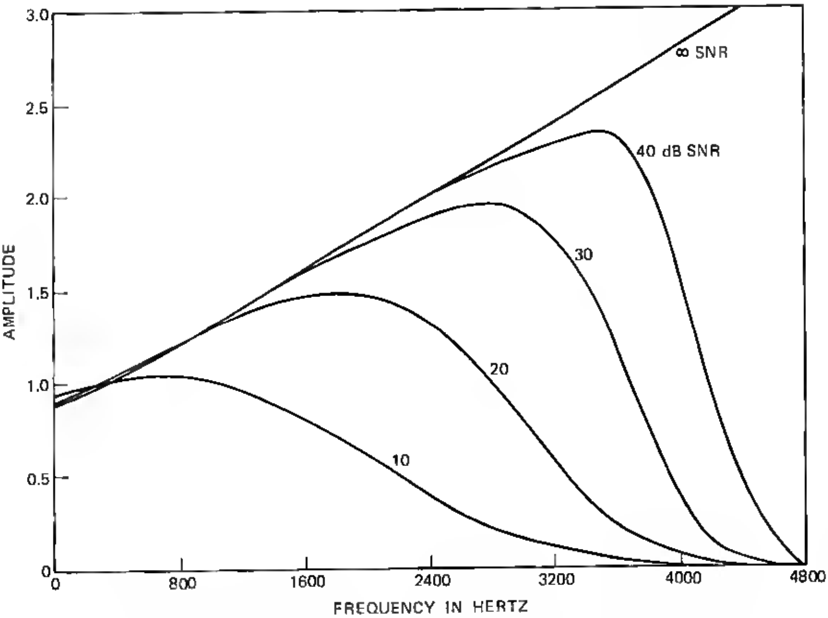

Fig. 7—Equalizer amplitude characteristics for various noise levels.

which is of the same form as the one-channel result [see (27)], as well as the filter obtained by Maurer and Franks (see Ref. 2). Clearly, for those frequencies where the noise is negligible (compared to the signal) the equalizer characteristic is $[1/\bar{a}(\omega)]e^{-j\bar{\theta}(\omega)}$, i.e., the best filter is one which *inverts* the average channel. Recalling the first question posed in the introduction, we can see that, apart from noise rejection (which occurs in the frequency range of small signal spectrum), the filter will invert the average channel when the variance of the ensemble is not appreciable. This is a useful rule-of-thumb for rapid (and approximate) compromise equalizer design.

IV. EXAMPLES USING THE 1964 CUSTOMER LOOP SURVEY

In this section we consider the design of a compromise equalizer for use over an ensemble of data customer loops. Some knowledge of the characteristics of this ensemble can be obtained from the 1964 Loop Survey.[5] This survey collected information about transmission parameters and channel makeup (e.g., gauges and lengths of sections which compromise the loop, locations of load coils and hridged taps, etc.) for

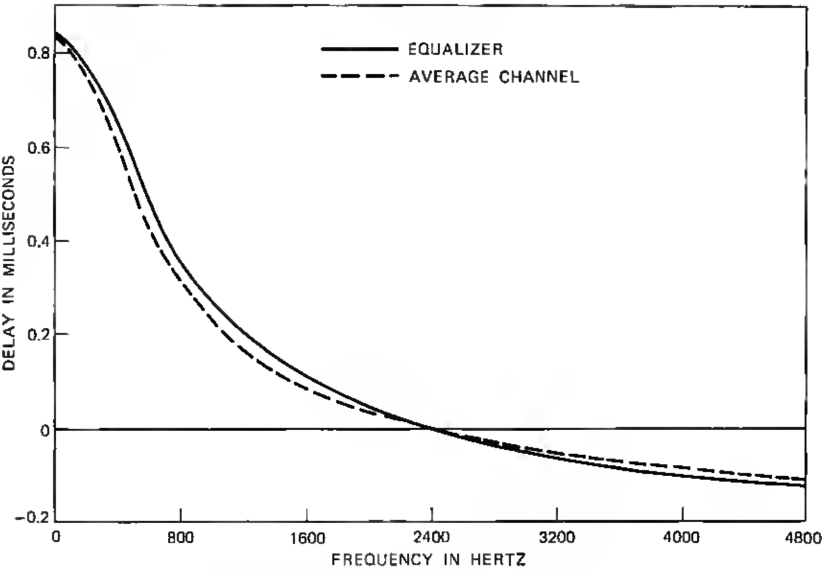

Fig. 8—Average channel delay (inverted) and equalizer delay (SNR = 30 dB).

a random selection of loops. For our purposes we extracted an ensemble of business loops (as opposed to residential loops) described in the survey. An existing computer program was used to obtain the frequency responses of the channels (with load coils and bridged taps removed) from the descriptions of the physical makeup of the channels. A recent study[6] shows that, at voiceband frequencies, the functions $\{e^{-\sqrt{i\omega\alpha}}\}_{\alpha>0}$ provide a good approximation of the loop transfer characteristics. The parameter α is a random variable (whose distribution can be approximated by the survey information), specifically $\alpha = RC\ell^2$, where R is the average series resistance per mile, C is the capacitance per mile, and ℓ is the loop length in miles. The average channel is depicted in Figs. 5a and 5b in terms of gain and delay.† These loop characteristics display appreciable amplitude variation (as well as delay variation) and the extent of these variations is indicated by the isolated circles.

To illustrate the design technique on these loop networks, we determined a compromise equalizer characteristic to be employed in a maxentropic 4.8-kb/s stream of randomly signed pulses. The basic

† The channel delays each include an estimate of B_{opt}.

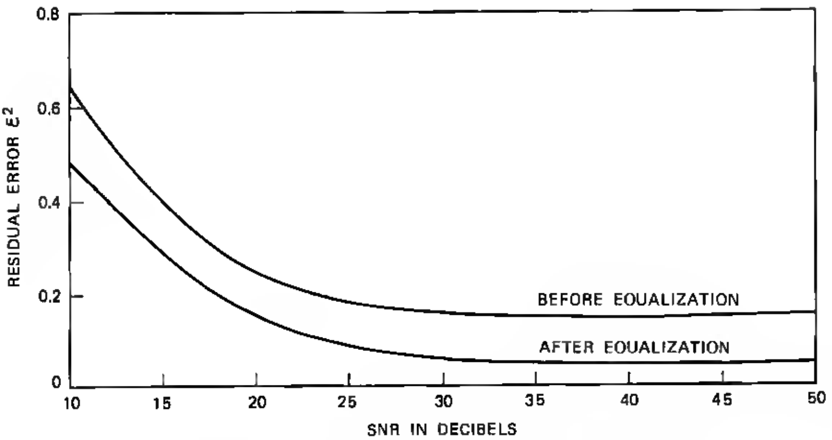

Fig. 9—Reduction of residual error as a result of equalization.

pulse was assumed to have a raised cosine spectrum, and the resulting power spectral density is shown in Fig. 6. It was assumed that only the "worst" 70 of the 143 customer loops required equalization. Figure 7 shows the resulting equalizer amplitudes for various noise levels. In this example, the channels were normalized to avoid the "whomper" effect mentioned in Section II. So when the signal-to-noise ratio is not infinite, the noise levels should be interpreted as if the lowest signal-to-noise ratio for the entire ensemble was assumed for all channels. Ohserve that, in the ahsence of noise, the equalizer tends to invert the average channel amplitude characteristic, and as the noise level is increased the equalizer tends to attenuate the high end of the spectrum where the signal power is lowest. Also note that, as the signal-to-noise ratio decreases, the solution tends to a filter matched to the raised cosine characteristic times a characteristic approximating the average channel. We ohserved that, for each of the noise levels, the compromise equalizer delay is close to the inverse delay of the average channel. The specific delay curve for the case of 30-dB SNR is provided in Fig. 8, which also shows the average channel delay (inverted). Finally, a direct computation shows that, for typical signal-to-noise ratios, the equalizer reduces the residual error hy 4.8 dB, and Fig. 9 displays this improvement in mean-square error as a function of signal-to-noise ratio. The improvement is greatest at high SNR since the equalizer is comhating only linear distortion rather than a comhination of linear distortion and noise.

VI. CONCLUSIONS

Using the second-order statistics of the channel ensemble, as well as the signal and noise spectra, the equalizer characteristics were easily computed by solving a matrix eigenvalue problem. Several interesting conclusions can be drawn concerning the properties of the equalizer:

(*i*) The equalizer amplitude is attenuated in those frequency regions where the signal-to-noise ratio or signal-to-channel-variance ratio is small.

(*ii*) When the channel ensemble has only delay distortion, amplitude as well as phase compensation is required.

(*iii*) The delay characteristics of the equalizer are rather insensitive to changes in noise level or (non-random) amplitude distortion.

(*iv*) When the channel ensemble has small variation about the average characteristic, a situation that commonly arises in practice, then the equalizer will invert the average channel.

REFERENCES

1. Bennett, W. R., and Davey, J. R., *Data Transmission*, New York: McGraw-Hill, 1965.
2. Maurer, R. E., and Franks, L. E., "Optimal Linear Processing of Randomly Distorted Signals," IEEE Trans. Circuit Theory, *17*, No. 1 (February 1970).
3. Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, New York: Interscience Publishers, 1966, Chapter 3.
4. Davenport, W. B., Jr., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, New York: McGraw-Hill, 1958.
5. Gresh, P. A., "Physical and Transmission Characteristics of Customer Loop Plant," B.S.T.J., *48*, No. 10 (December 1969), pp. 3337–3385.
6. Brophy, F., Foschini, G. J., and Smith, J. W., "Some Distortion Versus Speed Results for High-Speed Cable Transmission Systems," unpublished work.